Computers in Biology and Medicine 96 (2018) 8–23

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Investigating the contribution of distance-based features to automatic sleep stage classification

Ali Abdollahi Gharbali [a,b,*], Shirin Najdi [a,b], José Manuel Fonseca [a,b]

[a] CTS, Uninova, 2829-516, Caparica, Portugal
[b] Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa Campus da Caparica, Quinta da Torre, 2829-516, Monte de Caparica, Portugal

A R T I C L E I N F O

Keywords:
Polysomnography
Sleep stage classification
Feature extraction
Itakura
Itakura-saito
Feature selection
Distance-based features

A B S T R A C T

Objective: In this paper, the contribution of distance-based features to automatic sleep stage classification is investigated. The potency of these features is analyzed individually and in combination with 48 conventionally used features.

Methods: The distance-based set consists of 32 features extracted by calculating Itakura, Itakura-Saito and COSH distances of autoregressive and spectral coefficients of Electrocardiography (EEG) (C_3-A_2), Left EOG, Chin EMG and ECG signals. All the evaluations are performed on three feature sets: distance-based, conventional and total (combined distance based and conventional). Six ranking methods were used to find the top features with the highest discrimination ability in each set. The ranked feature lists were evaluated using *k*-Nearest Neighbor (*k*NN), Artificial Neural Network (ANN), and Decision-tree-based multi-SVM (DSVM) classifiers for five sleep stages including Wake, REM, N1, N2 and N3. Furthermore, the ability of distance-based and conventional features to discriminate between each pair of sleep stages was evaluated using *t*-test, a hypothesis testing method.

Results: Distance-based features occupied 25% of top-ranked features. Simulation results showed that using distance-based features together with conventional features can lead to an enhancement of accuracy. The best classification accuracy (85.5%) was achieved by DSVM classifier and 13 features selected by mRMR-MID and normalized with Min-Max method for total feature set, where two of them were from the distance-based feature set. The *t*-test results show that distance-based features outperform conventional features in discriminating between N1 and REM stages that is usually a challenge for classification systems.

Conclusion: Distance-based features have a positive contribution to sleep stage classification, including enhancement of accuracy and better REM-N1 discrimination ability.

Significance: The main motivation for this work was to evaluate new features to characterize each sleep stage in such a way that extracted features were more powerful than conventional features, to distinguish sleep stages from each other, and to improve classifiers accuracy.

1. Introduction

Sleep is fundamental for physical and mental health and occupies a significant part of human life. Therefore, the diagnosis of sleep-related disorders is of great importance for sleep research. Normal human sleep consists of two distinct stages with independent functions known as Non-Rapid Eye Movement (NREM) and Rapid Eye Movement (REM) sleep. In an ideal situation, NREM and REM states alternate regularly, each cycle lasting 90 min on average. NREM sleep accounts for 75–80% of sleep duration. According to the American Academy of Sleep Medicine (AASM) [1], NREM is subdivided into three stages: stage N1 or light sleep, stage N2 and stage N3 or Slow Wave Sleep (SWS). On the other hand, REM sleep accounts for 20–25% of sleep duration. The first REM state usually occurs 60–90 min after the onset of the NREM and lasts a few minutes [2]. To perform sleep analysis, specific physical and electrical activities of the body and brain are recorded. For this aim, a multiple-parametric test, called polysomnography (PSG) is usually used. During the PSG test, a number of biosignals including Electroencephalogram (EEG), Electro-oculogram (EOG), chin Electromyogram (EMG), leg Electromyogram (EMG), airflow signals, respiratory effort signals, oxygen saturation, body position, and electrocardiogram (ECG) are recorded in overnight sleep. The presence of skillful technicians and

* Corresponding author. CTS, Uninova, 2829-516 Caparica, Portugal.
E-mail addresses: a.gharbali@campus.fct.unl.pt (A.A. Gharbali), s.najdi@campus.fct.unl.pt (S. Najdi), jmrf@fct.unl.pt (J.M. Fonseca).

https://doi.org/10.1016/j.compbiomed.2018.03.001
Received 14 June 2017; Received in revised form 4 March 2018; Accepted 4 March 2018

physicians is necessary for assuring the quality of the recording and analysis. After the acquisition of the PSG, the data are scored by a technician according to a collection of rules set forth by AASM. According to these criteria, the scoring should be performed on 30 s, sequential epochs starting from the first sample of the data. For each stage, a number of recommended definitions are presented in AASM. These definitions mainly include EEG frequency and waveform, eye blinks and movements and EMG amplitude. EEG is divided into five frequency bands: slow wave activity (0.5–2 Hz), delta waves (0–3.99 Hz), theta waves (4–7.99 Hz), alpha waves (8–13 Hz) and beta waves (greater that 13 Hz). As an example, trains of 2–6 Hz saw-tooth waves with high frequency over the central head region in EEG, rapid eye movements in EOG, and low-chin EMG are typical indications of REM stage [1]. Manual scoring of sleep stages has several challenges and disadvantages. First of all, it is time consuming. It usually takes hours to score the PSG of a whole night of sleep. Second, the results of sleep scoring from two different practitioners are often not consistent. It has been reported that there is considerable inter-scorer variability (~20% disagreement) among scorers. Such differences are typically the result of rapid transitions between stages, which create ambiguous stages [3]. Moreover, with the immergence of at-home sleep monitoring systems, there is an urgent need for unsupervised methods that can efficiently score sleep data in a way that the results are medically reliable. Therefore, developing automatic sleep stage classification algorithms has been the focus of many researchers.

The common approach in automatic sleep stage classification, like any other pattern recognition process, includes feature extraction and classification steps. Features are extracted from a subset of raw PSG recordings containing only raw EEG data or EEG data together with other raw PSG signals acquired. Various types of features can be extracted from PSG recordings. Since EEG data in a transform domain reveals more useful information than in the time domain, usually a series of transformations such as Fourier transform [4–6], Short Time Fourier Transform (STFT) [7], Wavelet transform [8–12], Hilbert-Hung transform [13], and Empirical Mode Decomposition (EMD) [14–17] are applied. Because the information related to the sleep stages can be inferred from the EEG rhythms, the coefficients resulted from these transforms are used to extract EEG frequency bands. Moreover, these coefficients are often regarded as different representations of the PSG recordings. Therefore, several statistical and nonlinear features are extracted from them [18]. In Refs. [4,5] for instance, the mean of the absolute values, average power, and standard deviation are extracted from Discrete Wavelet Transform (DWT) coefficients of each EEG sub-band.

On the other hand, the popularity of the transform-based features does not imply that temporal or nonlinear ones extracted from PSG recordings are not useful in sleep stage classification. Temporal features have lower computational complexity and simulate the manual scoring process performed by the technician. Statistical parameters [4,14,19,20], autoregressive model parameters [21], Hjorth parameters [22,23], and features based on period analysis like zero-crossing rate [20,24] are among the most common temporal features extracted from PSG recordings, especially EEG.

Since the dynamic behavior of individual neurons is governed by threshold and saturation phenomena, nonlinearity is apparent in the brain's neural network. Moreover, the brain's ability to perform advanced cognitive tasks rejects the hypothesis of a completely stochastic brain. In addition to the EEG, other signals acquired from the body have neither a completely stochastic nature nor are stationary. Therefore, nonlinear signal processing techniques have also widely been used for characterizing sleep signals. Entropy estimators [23,25–27], fractal dimension [25,28], and Hurst exponent [29,30] are only some of the nonlinear features used in this area. In addition to the conventional features, looking at recent papers in sleep stage classification to understand the latest trends can be useful. There are few works in the area that present new features for enhancing the quality of the sleep staging systems. In Ref. [31], two new statistical features were applied to the single-channel EEG: Maximum-Minimum Distance (MMD) and EnergySis (*Esis*). To extract MMD, each EEG epoch is divided into sub-windows. MMD is defined as the Euclidean distance between the maximum and minimum points of the EEG waveform in the corresponding sub-window. MMD of each epoch is calculated by summing the MMDs of the sub-windows. Regarding the second feature, EnergySis, the basic idea is that the signal has energy and speed. This feature is calculated by multiplying the sum of the squared amplitude and velocity of each epoch.

Feature vector quality is an important factor for developing a reliable classification system. Features used in a specific machine-learning problem can perform reasonably well for other problems as well. Therefore, researchers often evaluate and explore the applicability of various features in different machine-learning areas. Kong et al. in Ref. [32] assumed that an EEG signal can be modeled as an autoregressive (AR) process and used Itakura distance to measure the similarity of the EEG signals. Itakura distance has, in fact, been found effective in distinguishing hypoxia and asphyxia. Later in 2004, Estrada et al. [33] used the Itakura distance for measuring the similarity of a baseline EEG epoch to the rest of the EEG in the context of sleep stage classification. In addition to the similarity of EEG signal with itself, in Refs. [34,35] it is demonstrated that the Itakura distance between EEG and EOG is also a useful similarity measure for sleep stage classification.

In the classification step, various types of classifiers have been used in the literature. Among them, Artificial Neural Network (ANN) [36–39], statistical classifiers such as Support Vector Machines (SVM) [4,9,40,41], instance-based classifiers like *k*-Nearest Neighbor (*k*NN) [40,42] are among the most popular classifiers. In principle, SVMs are designed for binary classification problems, but for the cases that discrimination among more than two classes is required, like sleep stage classification, a multi-class framework of SVM is developed. Several papers provide evidence for the high performance of SVM [4,9,40]. However, several practical challenges such as improving generalization and reducing computational complexity of these systems are still unsolved.

Considering the outstanding performance of Itakura and Itakura-Saito distances in sleep and speech signal processing and COSH distance in speech signal processing, this work extends our initial study on the distance-based features for sleep stage classification [23,45], where Itakura distance outperformed conventional features in classifying the sleep data from the Physionet database [46]. Since the works presented so far use distance-based features in a restricted manner (only the performance of the limited variations are tested), in the current work we aim to extensively evaluate the performance of distance-based features together with conventional features in automatic sleep stage classification. These distance-based features are extracted by calculating Itakura, Itakura-Saito and COSH distances of autoregressive and spectral coefficients of EEG, EMG, EOG and ECG signals.

The following contributions to improve automatic sleep stage classification are presented:

a) Evaluating the potency of distance-based features for sleep stage classification,
b) Comparing the performance of a distance-based feature set with conventionally used features,
c) Assessing the effect of feature normalization on classification results,
d) Utilizing the Vikor method for finding the optimum number of features considering classification accuracy,
e) Analyzing discrimination ability of top features selected by different feature ranking methods, including conventional and distance-based, using statistical hypothesis testing method.

The rest of the paper is structured as follows: Section 2 explains the database used in this work. Section 3 provides a detailed description of the study framework used in this paper. In section 4, the performance of various parts of the framework is assessed and the results are presented. Section 5 finalizes the paper with the conclusions and direction of the future work.